Stereo matching algorithm based on combined matching cost computation and edge preserving filters

Madiha Zahari[1,2], Rostam Affendi Hamzah[2], Nurulfajar Abd Manap[1], Adi Irwan Herman[3]
[1]Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2]Faculty of Electrical and Electronic Engineering Technology, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[3]Product and Test Engineering Department, Texas Instruments, Batu Berendam, Melaka, Malaysia

Article Info

Article history:

Received Aug 26, 2021
Revised Mar 22, 2022
Accepted Apr 7, 2022

Keywords:

Census Transform
Disparity Map
Matching Cost
SAD
Stereo Matching

ABSTRACT

The stereo matching process is one of the key areas that impact the stereo vision technologies which are commonly used in the application of three-dimensional reconstructions. The accuracy of the depth information used in three-dimensional reconstruction is directly proportional to the accuracy of the disparity obtained from stereo matching. The challenging issue in the stereo matching process is to determine the accurate corresponding point between the left image and right image, especially for image pairs that have different exposure such as different illumination and image pair with less texture region. In order to increase the accuracy of disparity value, a new stereo matching algorithm is proposed based on the combination of Sum of absolute different and census transform at matching cost computation. guided filter was used in the matching cost aggregation in order to remove noise and preserve the edge of the image. In the optimization step, the winner take all strategy is used to select the minimum matching cost. Finally, a median filter is applied to the initial disparity map for refinement purposes. The experimental results show that the algorithm is effective in reducing the error and improving the accuracy of the disparity map in different illumination regions, less textured regions and different environmental exposure.

Corresponding Author:

Madiha Zahari
Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, 76100, Melaka, Malaysia
Email: madiha@utem.edu.my

1. INTRODUCTION

Computer vision is a significant research topic in the computer system due to increased usage of the autonomous navigation system, industrial monitoring, unmanned driving, virtual reality, image-based rendering and vision-based object handling. All the usage required accurate vision system in order to ensure the system works as designed. One of the aimed of computer vision is to provide an accurate image information and to reform its properties in terms of its shape, illumination, and colour distribution. Stereo vision is one of the main areas of research in the computer vision field. It is the technology inspired by human eyes where this system consists of two cameras captured a scene simultaneously by the cameras and then processed the images to get the information about the distance between the object to the images which known as depth value. Stereo matching is the common method used to determine the depth information about the stereo images. This method is used in order to determine the disparity value of all pixels in stereo images [1]. By calculating the differences of pixel value at the left image and the right image at two corresponding points, the disparity value of both images is determined. The pre-processing algorithm is applied to the images for image rectification in order to find the disparity value easily [2].

The process of determining the depth value of the stereo images are depending on how the algorithm is designed to compute the disparity value. The process can be classified into two major categories which is global method and local method [3]. By using global method, the disparity values are computed by optimizing a general energy function for all pixels of an image [4]. This method gives a higher accuracy, but it also required high computational complexity [5]. On the other hand, local method which calculated the cost volume in a limited area within the image gives lower accuracy in stereo matching compared to global method. Because of less complexity in computational steps, local methods are used in most real time application of stereo system [6]. Scharstein *et al*. [3] proposed four steps of the stereo matching algorithm which consist of matching cost computation, cost aggregation, disparity optimization and disparity refinement. The challenging problem in stereo matching algorithm is to increase the accuracy of the depth value. The correct matching at low texture region become one of the challenging in getting an accurate disparity map since the possibility of matching with invalid correspondence point is high. The illumination difference between the image pair also will affect the accuracy of the disparity map generate by the matching algorithm. Due to this issue, this paper proposed a framework of stereo matching algorithm which aim to improve the accuracy of the depth value of the less textured region and on the image with illumination differences. In recent years, there are varieties of stereo matching algorithm had been proposed by researchers to increase the accuracy of the stereo matching algorithm. Zhang *et al*. [7] had proposed cross-based stereo matching which gave an accurate disparity estimation compared with global method. Yang *et al*. [8] proposed a simple local matching algorithm which efficient to be implemented.

The most common matching cost computation such as the sum of absolute different, sum of squared difference (SSD) and normalized cross correlation (NCC) are the traditional similarity measure function for stereo matching using the block matching method. However, this method is very sensitive to the amplitude distortion which leading to low accuracy [9]. A non-parametric local transform had been introduced which consist of census transform (CT) and rank transform (RT) [10]. Both methods are more resistant to radiometric distortion because they depend on the relative order of pixel intensity instead of the intensity values themselves [11]. Therefore, both types of non-parametric local transforms able to cope with the matching uncertainties well for image regions with similar colors, while SAD and SSD may cope well with the image regions with similar local structures [12]. Due to this factor, researchers started to introduce the combination of matching cost computation. The combination of CT with other cost matching method gain the researcher attention from the performance of the combined matching cost showed better results. Work done by Lee *et al*.[11] proposed combination of CT with gradient distance as cost computation method. Zhu *et al*. [13] also used combination of CT with gradient as matching cost method. Wang *et al*. [14] used combination of CT and absolute difference as cost computation method. The combination of SAD algorithm and gradient matching in cost computation also produced better accuracy of disparity value [15].

Cost aggregation is the second step in stereo matching that directly influence the overall efficiency and the accuracy of the matching algorithm. In most local matching algorithm, filtering based cost aggregation methods are commonly adopted in this process where the process of filtering the matching cost volume happened. The simplest approach that can be used such as the Box filter and the Gaussian filter where the aggregation is only in the fixed window size [13]. However, these filtering techniques yield a bad performance of the disparity map with fatten edges. In local methods, guided filter (GF) and bilateral filter (BF) become popular edge-preserving filters used as a cost aggregation method since both methods able to produce good quality of disparity map with fine edges and also produce disparity map with better accuracy [16]. GF has gained better performance and efficiency compared with BF since the complexity of GF algorithm is lower compared to BF. GF become a popular approach used by researchers in order to develop an accurate stereo matching algorithm [17]. The study done by Hosni *et al*. [18] showed that GF is robust to be applied in the stereo matching area. There are various cost aggregation method had been applied in stereo matching algorithm based on that study. Weighted GF had been proposed by Kong *et al*. [19] yield a better performance of GF. Cost volume filtering method which proposed by Hong and Kim [20] resulted an improved disparity map. The adaptive weight of the local variable is used to control the linear coefficient based on the local texture features. Wu *et al*.[21] proposed to combine GF with minimum spanning tree (MST) filter which able to increase the robustness of highly textured and texture less regions. Hamzah *et al*. [22] introduced an adaptive support weight based on iterative GF. Moreover, Zhu and Chang [23] proposed an innovative weighted-combination scheme of GF model which improved matching cost volume. In this paper GF is adopted as cost aggregation method which aimed to produce a disparity map with fine edge and able to improve the accuracy of the results.

This paper proposed a stereo matching algorithm that uses a combination of two measurement methods in cost computation steps which are SAD and CT. The SAD cost computation which is very sensitive to the amplitude distortion and illumination difference and by combining with CT cost function, higher accuracy and robust cost function was developed. GF is selected as a filter-based cost aggregation method that aims to improve the quality of the depth map at the edges and the discontinuity zone. Then the process of optimization used the winner-take-all (WTA) strategy. At this point, each valid pixel is absorbed based on the lowest aggregated corresponding value. However, certain invalid pixels are always present in the less texture areas. The final stages consist of post-processing works which used weighted median (WM) filter in order to obtain a more accurate disparity map.

2. RESEARCH METHOD

The SAD method has been used in order to measure the intensity difference between left and right images. The measurement of SAD has been implemented using (1):

$$SAD(p,d) = \sum\nolimits_{i\in\{R,G,B\}} \left(I_l^i(p) - I_r^i(p-d)\right) \tag{1}$$

where the pixel at coordinates $(x;y)$ denote by p, i means the RGB channels number, d is the disparity value, I_l represents the left image and I_r represents the right image. CT process maps the neighbouring surrounding pixel to a bit string which can denote the intensity value of the neighbouring pixel [10]. The process of CT are based on the (2):

$$CT(p) = \otimes_{q\in w_{CT}}\, cen(p,q) \tag{2}$$

where p and q represent the target pixel and neighboring pixels respectively and $\otimes$ refer to the eXclusive OR (XOR) process of the bit value for $cen(p,q)$ with w_{CT}, window size and $cen(p,q)$ represent the binary function with the condition as given by (3):

$$cen(p,q) = \begin{cases} 1, I(p) \geq I(q) \\ 0, otherwise \end{cases} \tag{3}$$

where $I(p)$ and $I(q)$ are the target pixel and neighboring pixels values respectively. By using the CT process implemented as (2), the cost volume at each corresponding pixel is calculated by calculating the different between two bit strings which is given by (4):

$$CT'(p,d) = HD(CT_l(p) - CT_r(p-d)) \tag{4}$$

where CT_l is the bit string obtained from the CT process of the left images and CT_r is the bit string obtained from the CT process of the right images. The integrated of the two matching cost is based on the normalized cost function proposed by [11], which represent in (5):

$$M_c(p,d) = 2 - exp\left(-SAD(p,d)\right) - exp(-CT'(p,d)) \tag{5}$$

In this work, the initial matching cost was aggregated using guided filter (GF) in order to ensure the noise can be removed from the cost and the edge of can be preserved [13]. This filtering method used a reference image as a guidance during filtering process and for this work, the left grayscale image is used as the reference image. The filter kernel of the guided filter [17] is defined as (6):

$$GF_{p,q}(I_n) = \frac{1}{|s|^2}\sum\nolimits_{q\in w_k}\left(1 + \frac{(I_{p,n-1}-\mu_{k,n-1})(I_{p,n-1}-\mu_{k,n-1})}{\sigma_{k,n-1}^2+\varepsilon}\right) \tag{6}$$

where I_n represents the reference grayscale image, p is the coordinates for target pixel, $(x;\, y)$. w_k is the window size with the size of $r\ x\ r$ pixels. s is the sum of pixels in the w_k, q is the neighboring pixel and k is the center pixel. The variance and the average of the intensity values in the reference image represent by σ and μ. The ε is the smoothness term control element. At this step the matching cost is aggregated and the total cost, $CA(p,d)$ is define as (7):

$$CA(p,d) = M_c(p,d)GF_{p,q}(I_n) \tag{7}$$

After completing the matching aggregation, winner-takes-all (WTA) strategy was implemented at this step in order to select the minimum matching cost as the initial disparity value [3]. The initial disparity, d_i is define as (8):

$$d_i = arg\, min_{d \in R}\, CA(p, d) \quad (8)$$

where R is all the possible disparity value. Lastly, the weighted median filter of is used to refine the map of disparities [22]. Let D indicate the resulting disparity map after adjusting the normal base plane. The final disparity map,D^F is refined as (9).

$$D^F(p) = \underset{q \in \Omega}{med}\, \{ d_i \} \quad (9)$$

3. EXPERIMENTAL RESULTS AND DISCUSSION

The experiment was conducted to assess and determine the performance of the proposed algorithm from various perspectives. The results were evaluated using Middlebury dataset 2014 which consists 15 training images [24]. The absolute disparity error in non-occluded area, *nonocc* and entire image area, *all pixels* had been measured using online Middlebury evaluation system. The performance comparison between the proposed matching cost and other three matching cost was evaluated. Tabel 1 shows the average error by using four different matching cost computation method with 15 Middlebury dataset training images. Figure 1(a) shows all the data for absolute average error (*nonocc*) for all training images and Figure 1(b) shows data for all pixel absolute average error (*all*). The combination of SAD and CT give the lowest percentage of error which 8.37% and 17.8% of average error for *nonocc* and *all* pixels respectively. While CT gives 9.45% for *nonocc* and 19% for all pixels. Combination of CT and Gradient matching cost give 9.16% and 18.6% of average error for *nonocc* and all and the matching cost using gradient methods give 9.97% and 19.4%. Based on this data, it shows that the combination of SAD and CT gives better accuracy compared with other three matching cost computation method.

In cost aggregation step, the evaluation was done by comparing the performance of the GF with BF and BOX filter using Middlebury training dataset. Figure 1 shows the results of the comparison between GF and other two CA methods based on the average absolute error and percentage of bad pixels. The results show that the GF provided higher accuracy compared to BOX. However the performance of GF compared to BF is almost the same, but still gave better accuracy. Figure 2 shows the example of the disparity map produced from the evaluated algorithm. Smooth disparity map obtained when GF and BF used as CA methods. The edges are well preserved when using GF compared with the disparity map obtained by using BOX which produced fatten edge. The results shows that the combination of SAD and CT at matching cost is worked better by combining with the GF compared to other aggregation method.

Figure 3 shows the samples of left and right input images of ArtL and PianoL taken from Middlebury dataset. Both sets of images have different illumination between left image and right images which cause very challenging to be matched since the pixel amplitude at the same corresponding point is totally different. However, the proposed algorithm manages to discover the corresponding point. The overall results show that the proposed algorithm can generate a disparity map which is better and competitive with the other established framework. Figure 3 also shows the comparison of the disparity map produced by using proposed algorithm and other framework. The image of PianoL also has an area on the floor, which considered as textureless region and the results in Figure 3 shows that the smooth disparity map obtained at that area by using the proposed algorithm compared to other framework. Based on the disparity map of PianoL it can see that the proposed algorithm also has the ability to produce a minimum error at less texture region with different illumination.

Table 2 shows the quantitative evaluation of the absolute error for the Middlebury Dataset based on proposed algorithm compared with other frameworks. R-NCC, BSM and proposed method used traditional stereo matching algorithm framework while DSGCA, DoGGuided and DF are using artificial intelligence frameworks. The overall result shows that the proposed algorithm is the lowest percentage of average disparity error for non-occluded region and all pixels. The results also show that the absolute error produced by using proposed algorithm is the lowest compared to other framework for images with less texture such as Adiron, Recycle and Playroom. The proposed algorithm also gives the best accuracy in the image with different exposure such as MotorcycleE. The overall results also show that the proposed method gives better accuracy compared with another method, including framework using an artificial intelligence-based framework.

Table 1. Comparison of percentage of average error between four types of matching cost computation for non-occluded and all pixels region.

Type of MCC	SAD +CT		CT		GRAD		CT+GRAD	
Middlebury	*nonocc*	*all*	*nonocc*	*all*	*nonocc*	*all*	*nonocc*	*all*
Avg	8.37	17.8	9.45	19	9.97	19.4	9.16	18.6
Adiron	2.98	7.41	3.14	7.84	15	19.6	9.74	14.3
ArtL	5.46	22.2	6.19	23.4	7.82	24.7	8.57	25.3
Jadepl	16.9	43.5	17.5	44.1	18.7	44.9	16.7	41.7
Motor	3.47	11.1	3.78	11.5	3.8	11.3	3.85	11.6
MotorE	3.37	11	3.67	11.4	8.46	16.6	5.91	13.9
Piano	5.82	10.6	6.68	11.6	7.33	11.9	7.56	12.3
PianoL	20.3	24.1	20.8	24.9	34.1	38.1	30.4	34.4
Pipes	7.47	20.1	7.82	20.7	7.45	20.4	7.78	20.6
Playrm	6.23	23	6.83	23.9	11.7	27.2	9.01	24.1
Playt	31.7	36.3	35.8	40.2	18.2	24.2	23.3	29.3
PlaytP	7.44	14.2	12.4	19	3.8	10.9	5.61	13.2
Recyc	3.79	7.95	4.16	8.36	4.48	8.06	4.01	8.48
Shelvs	12.7	15.7	13.3	16.3	16.2	18.9	14.3	17.4
Teddy	2.72	12.4	3.07	13	3.02	11.5	2.82	11.8
Vintge	19.5	25.7	22.7	28.7	9.42	16.7	7	13.7

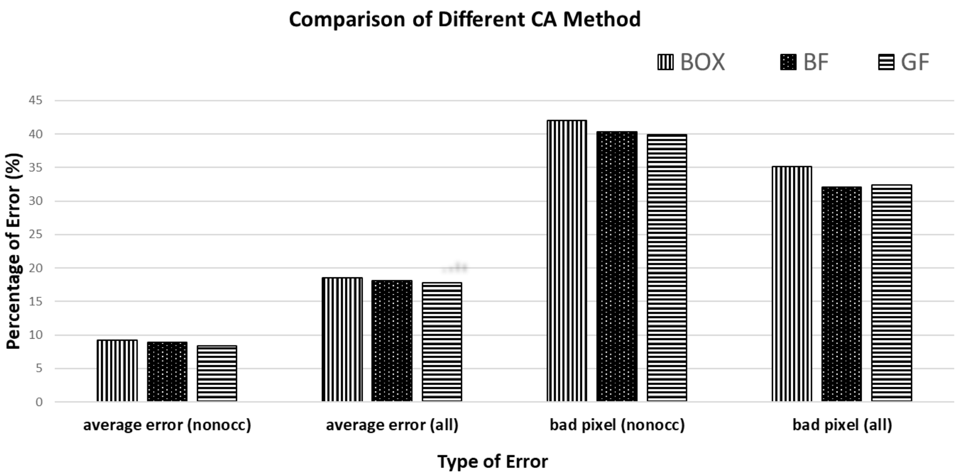

Figure 1. Percentage of average absolute error and average of bad pixels for non-occluded and all pixels by using GF, BF and BOX as cost aggregation method

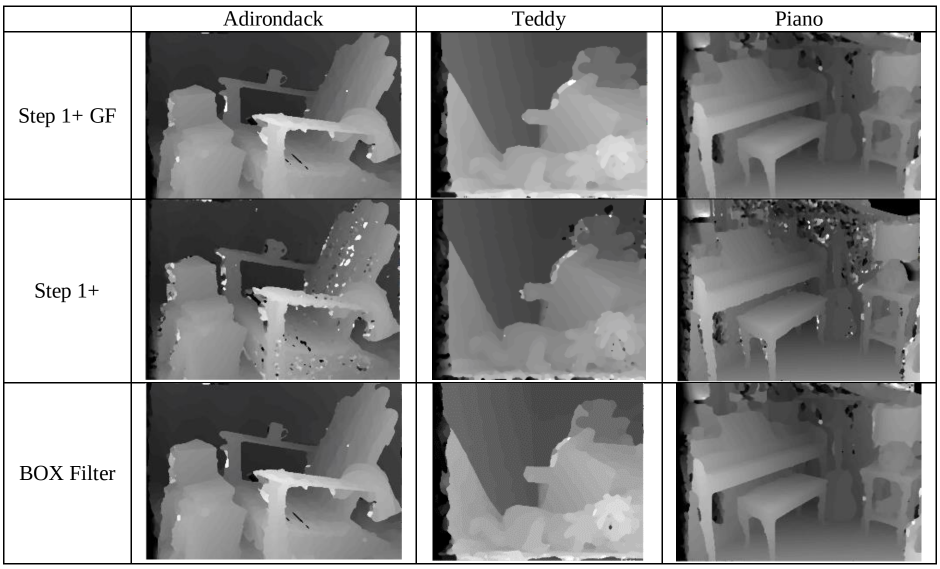

Figure 2. Example of disparity map using GF, and BOX Filter

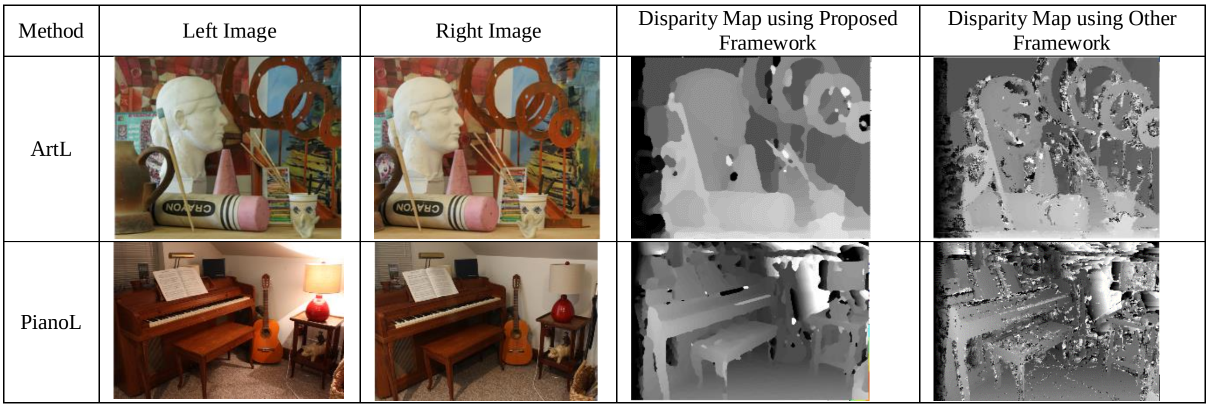

Figure 3. Sample of left and right input image from middlebury dataset and disparity map by using proposed algorithm and other framework for input image with different illumination

Table 2. The results of the quantitative evaluation of absolute error for all pixels and non occluded region using Middlebury dataset

Method	R-NCC [25] %		DSGCA [26] %		BSM [27] %		DoGGuided [28] %		MANE [29] %		DF [30] %		Proposed %	
	non occ	all	non occ	all	non occ	all	non occ	all	non occ	all	non occ	all	non occ	all
Adiron	20.5	21.2	3.25	7.68	7.27	12.7	15.2	20.1	6.58	11.6	13.2	14.1	2.98	7.41
ArtL	10	12.5	5.95	21.7	11.4	28.7	9.57	28	5.81	22.9	16.4	18.2	5.46	22.2
Jadepl	67.2	91	18.9	45	30.5	58.7	27.1	56.5	20.7	45.9	77.8	103	16.9	43.5
Motor	9.59	11.5	3.6	10.6	6.67	14.8	5.64	13.8	4.52	12.4	11.2	13.2	3.47	11.1
MotorE	10.6	12.7	3.41	10.4	6.52	14.7	8.31	16.8	4.31	12.3	10.7	12.7	3.37	11
Piano	9.12	9.59	7.17	11.5	10.8	16	8.09	13.4	10.6	15.1	10.5	11.1	5.82	10.6
PianoL	15.8	15.8	21.1	24.5	32.1	35.8	32.4	37.3	20.9	24.7	26.4	26.4	20.3	24.1
Pipes	21.8	27.9	7.23	19.9	10.5	24.5	9.67	23.8	8.62	22.3	16.1	22.5	7.47	20.1
Playrm	29	30	9.36	24.6	12.5	29.4	14	30.3	15	31.1	19.6	20.9	6.23	23
Playt	18	17.5	29.4	34.5	24.4	31	24.5	30.8	34.7	39.9	13.3	13.9	31.7	36.3
PlaytP	13.1	13	7.94	14.8	12.8	20.2	5.32	13	10.5	17.3	14.8	16.3	7.44	14.2
Recyc	22.3	22.2	3.8	7.56	7.42	12.1	5.56	9.13	5.5	9.67	16.2	16.8	3.79	7.95
Shelvs	11.5	11.7	14.7	17.3	16.4	19.2	16.2	19	20.2	22.5	11.1	11.5	12.7	15.7
Teddy	4.13	4.81	3.51	12.2	4.88	14.3	4.15	13.4	3.12	12.5	5.04	6.16	2.72	12.4
Vintage	44.3	45.1	39.7	43.8	32.8	39.3	15	23.6	46.5	51	24.9	26.8	19.5	25.7
Average	19.8	22.9	9.75	18.7	13.4	23.5	12	22.3	11.9	21.3	19.2	22.7	8.37	17.8

4. CONCLUSION

In this work, we proposed a local stereo matching algorithm which consists of the combination of SAD and CT in matching cost computation, GF at cost aggregation, WTA at disparity optimization and Median Filter (MF) at post processing. These combination steps are able to reduce the average absolute error from 10% to 19% compared with other matching algorithms. Besides, by using proposed matching cost, lowest absolute error is obtained compared with other methods including the disparity map of image pairs with different illuminations and exposure. The accuracy of the image with less texture also increased by using the proposed algorithm. Based on the results, it summarizes that the local stereo matching algorithm using the proposed method is able to reduce the matching error in less texture region and at different illumination region and also well preserved the edge of the image.

ACKNOWLEDGEMENTS

This work was supported by the Universiti Teknikal Malaysia Melaka with the grant number FRGS/1/2020/TK0/UTEM/02/12.

REFERENCES

[1] Y. J. Chang and Y. S. Ho, “Disparity map enhancement in pixel based stereo matching method using distance transform,” *Journal of Visual Communication and Image Representation*, vol. 40, pp. 118–127, Oct. 2016, doi: 10.1016/j.jvcir.2016.06.017.

[2] Y. S. Kang and Y. S. Ho, "An efficient image rectification method for parallel multi-camera arrangement," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1041–1048, Aug. 2011, doi: 10.1109/TCE.2011.6018853.

[3] D. Scharstein, R. Szeliski, and R. Zabih, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," in *Proceedings - IEEE Workshop on Stereo and Multi-Baseline Vision*, SMBV 2001, 2001, pp. 131–140, doi: 10.1109/SMBV.2001.988771.

[4] M. Dehnavi and M. Eshghi, "Cost and power efficient FPGA based stereo vision system using directional graph transform," *Journal of Visual Communication and Image Representation*, vol. 56, pp. 106–115, Oct. 2018, doi: 10.1016/j.jvcir.2018.09.002.

[5] L. Shi, F. Guo, W. Gao, and Z. Hu, "Stereo matching with adaptive support-weight correlation and graph cuts," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, Oct. 2010, pp. 3575–3579, doi: 10.1109/ICSMC.2010.5642342.

[6] Y. Hou, C. Liu, B. An, and Y. Liu, "Stereo matching algorithm based on improved Census transform and texture filtering," *Optik*, vol. 249, p. 168186, Jan. 2022, doi: 10.1016/j.ijleo.2021.168186.

[7] K. Zhang, J. Lu, and G. Lafruit, "Cross-based local stereo matching using orthogonal integral images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 1073–1079, Jul. 2009, doi: 10.1109/TCSVT.2009.2020478.

[8] D. W. Yang, L. C. Chu, C. W. Chen, J. M. Gan, J. Wang, and M. Der Shieh, "Low complexity stereo matching algorithm using adaptive sized square window," in *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test, VLSI-DAT 2014*, Apr. 2014, pp. 1–4, doi: 10.1109/VLSI-DAT.2014.6834897.

[9] Y. Chai and X. Cao, "Stereo matching algorithm based on joint matching cost and adaptive window," in *Proceedings of 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference, IAEAC 2018*, Oct. 2018, pp. 442–446, doi: 10.1109/IAEAC.2018.8577495.

[10] R. Zabih and J. Woodfill, "Non-parametric local transforms for computing visual correspondence," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 801 LNCS, pp. 151–158, 1994, doi: 10.1007/bfb0028345.

[11] Z. Lee, J. Juang, and T. Q. Nguyen, "Local disparity estimation with three-moded cross census and advanced support weight," *IEEE Transactions on Multimedia*, vol. 15, no. 8, pp. 1855–1864, Dec. 2013, doi: 10.1109/TMM.2013.2270456.

[12] H. Hirschmuller and D. Scharstein, "Evaluation of stereo matching costs on images with radiometric differences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1582–1599, Sep. 2009, doi: 10.1109/TPAMI.2008.221.

[13] S. Zhu, Z. Wang, X. Zhang, and Y. Li, "Edge-preserving guided filtering based cost aggregation for stereo matching," *Journal of Visual Communication and Image Representation*, vol. 39, pp. 107–119, Aug. 2016, doi: 10.1016/j.jvcir.2016.05.012.

[14] W. Wang, J. Yan, N. Xu, Y. Wang, and F. H. Hsu, "Real-time high-quality stereo vision system in FPGA," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 10, pp. 1696–1708, Oct. 2015, doi: 10.1109/TCSVT.2015.2397196.

[15] R. A. Hamzah, M. G. Y. Wei, and N. S. N. Anwar, "Development of stereo matching algorithm based on sum of absolute RGB color differences and gradient matching," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 3, pp. 2375–2382, Jun. 2020, doi: 10.11591/ijece.v10i3.pp2375-2382.

[16] Q. Yang, P. Ji, D. Li, S. Yao, and M. Zhang, "Fast stereo matching using adaptive guided filtering," *Image and Vision Computing*, vol. 32, no. 3, pp. 202–211, Mar. 2014, doi: 10.1016/j.imavis.2014.01.001.

[17] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013, doi: 10.1109/TPAMI.2012.213.

[18] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, Feb. 2013, doi: 10.1109/TPAMI.2012.156.

[19] L. Kong, J. Zhu, and S. Ying, "Local stereo matching using adaptive cross-region-based guided image filtering with orthogonal weights," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–20, May 2021, doi: 10.1155/2021/5556990.

[20] G. S. Hong and B. G. Kim, "A local stereo matching algorithm based on weighted guided image filtering for improving the generation of depth range images," *Displays*, vol. 49, pp. 80–87, Sep. 2017, doi: 10.1016/j.displa.2017.07.006.

[21] W. Wu, H. Zhu, S. Yu, and J. Shi, "Stereo matching with fusing adaptive support weights," *IEEE Access*, vol. 7, pp. 61960–61974, 2019, doi: 10.1109/ACCESS.2019.2916035.

[22] R. A. Hamzah, A. F. Kadmin, M. S. Hamid, S. F. A. Ghani, and H. Ibrahim, "Improvement of stereo matching algorithm for 3D surface reconstruction," *Signal Processing: Image Communication*, vol. 65, pp. 165–172, Jul. 2018, doi: 10.1016/j.image.2018.04.001.

[23] C. Zhu and Y. Z. Chang, "Hierarchical guided-image-filtering for efficient stereo matching," *Applied Sciences (Switzerland)*, vol. 9, no. 15, p. 3122, Aug. 2019, doi: 10.3390/app9153122.

[24] D. Scharstein *et al.*, "High-resolution stereo datasets with subpixel-accurate ground truth," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8753, pp. 31–42, 2014, doi: 10.1007/978-3-319-11752-2_3.

[25] Y. Li and S. Fang, "Removal-based multi-view stereo using a window-based matching method," *Optik*, vol. 178, pp. 1318–1336, Feb. 2019, doi: 10.1016/j.ijleo.2018.10.126.

[26] Williem and I. K. Park, "Deep self-guided cost aggregation for stereo matching," *Pattern Recognition Letters*, vol. 112, pp. 168–175, Sep. 2018, doi: 10.1016/j.patrec.2018.07.010.

[27] K. Zhang, J. Li, Y. Li, W. Hu, L. Sun, and S. Yang, "Binary stereo matching," in *Proceedings - International Conference on Pattern Recognition*, 2012, pp. 356–359.

[28] M. Kitagawa, I. Shimizu, and R. Sara, "High accuracy local stereo matching using DoG scale map," in *Proceedings of the 15th IAPR International Conference on Machine Vision Applications, MVA 2017*, May 2017, pp. 258–261, doi: 10.23919/MVA.2017.7986850.

[29] H. D. Vázquez-Delgado *et al.*, "Real-time multi-window stereo matching algorithm with fuzzy logic," *IET Computer Vision*, vol. 15, no. 3, pp. 208–223, Apr. 2021, doi: 10.1049/cvi2.12031.

[30] W. Mao and M. Gong, "Disparity filtering with 3D convolutional neural networks," in *Proceedings - 2018 15th Conference on Computer and Robot Vision, CRV 2018*, May 2018, pp. 246–253, doi: 10.1109/CRV.2018.00042.

BIOGRAPHIES OF AUTHORS

Madiha Zahari received the Bachelor degree in Electrical and Electronic Engineering from University Technology of Petronas in 2006 and Master of Engineering in Industrial Electronic and Control from University of Malaya in 2013. She had experienced working in semiconductor industries from 2006 until 2009. She is currently a lecturer in Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik, Universiti Teknikal Malaysia Melaka. She also one of the MIET Chartered Engineer. Her current research interests are in image processing and stereo vision. She can be contacted at email: madiha@utem.edu.my.

Rostam Affendi Hamzah graduated from Universiti Teknologi Malaysia where he received his B.Eng majoring in Electronic Engineering. Then he received his M. Sc. majoring in Electronic System Design Engineering and PhD majoring in Electronic Imaging from the Universiti Sains Malaysia. Currently he is a lecturer in the Universiti Teknikal Malaysia Melaka teaching digital electronics, digital image processing and embedded system. He can be contacted at email: rostamaffendi@utem.edu.my.

Nurulfajar Abd Manap received a Master of Electrical Engineering in image processing (2002, Universiti Teknologi Malaysia) and Bachelor of Electrical Engineering (2000, Universiti Teknologi Malaysia). He obtained his PhD (Image and Video Processing) at the University of Strathclyde, Glasgow, in 2012. He teaches various electronics engineering subjects and courses in his affiliation for undergraduate and postgraduate levels. His subjects include Advanced Digital Signal Processing, Computer Vision and Pattern Recognition, Data Structures, Multimedia Technology and Applications and Distributed and High-Performance Computing. He also one of the MIET Chartered Engineer and Apple Distinguished Educators (ADE). His research interests are 3D image processing, stereo vision and video processing). He can be contacted at email: nurulfajar@utem.edu.my.

Prof. Dr. Mohd Ali Hassan graduated in 2015 with a Bachelor Degree in Computer Engineering Technology (Computer System) from the Universiti Teknikal Malaysia Melaka. Currently, he works with Texas Instrument more than 5 years with his current research interests are computer engineering related field of studies. He can be contacted at email: adiirwanherman@gmail.com.